

## 案例实战：陌生人社交APP的MySQL索引设计实战（二）

---

今天我们继续分析这个社交APP的复杂用户搜索功能场景下的索引设计案例，上次我们讲到，在我们的这个场景里，SQL中会包含where、order by和limit几个语句，而且实际场景中，往往where和order by是没法都用到索引的，这是第一个我们要注意的问题。

今天我们来分析第二个问题，就是在where和order by出现索引设计冲突，鱼与熊掌不可兼得的时候，到底是针对where去设计索引，还是针对order by设计索引？到底是让where去用上索引，还是让order by用上索引？

其实这个问题的本质就是说，你是要让where语句先基于联合索引去进行一个筛选，筛选出来一部分用户指定的数据，接着再把数据加载到内存或者是基于临时磁盘文件去进行指定条件的排序，最后用limit语句拿到一页数据吗？

还是说要让order by语句按照你的索引的顺序去找，找的过程中基于where里的条件筛选出来指定的数据，然后再根据limit语句拿出来一页数据？

说实话，一般这种时候往往都是让where条件去使用索引来快速筛选出来一部分指定的数据，接着再进行排序，最后针对排序后的数据拿出来一页数据。

因为基于索引进行where筛选往往可以最快速度筛选出你要的少部分数据，如果筛选出来的数据量不是太大的话，那么后续排序和分页的成本往往不会太大！

好，那么假设我们打定主意要针对where条件去设计索引的话，此时又要犯难了，因为这个时候你要去考虑，用户在搜索潜在好友的时候，一般会上哪些条件呢？我们到底要把哪些字段包含到索引里去？到底在联合索引里，字段的顺序要如何排列呢？

其实开门见山要告诉大家的一点就是，我们首先应该在联合索引里包含省份、城市、性别，这三个字段！

因为这三个字段都是在搜索里几乎必定包含的三个字段，假设你要搜索潜在好友，那么必定是会搜索跟你同一个地方的，然后搜索某个性别的其他用户，这几个条件在APP里完全可以做成必选项，用户也几乎必定会指定。

但是此时有人就会说了，之前不是说过么，基数太低的字段最好别放到索引里去，那省份、城市和性别，都是基数非常小的几个字段，可选的值就那么几个，为什么要放到索引里去？

这是个好问题，但是规则是死的，人是活的。

假设你就因为省份、城市和性别几个字段的基数太小了，此时就不把他们几个包含到联合索引里去，那么你实际查询的时候都要基于这几个字段去搜索，此时你就只能把这几个字段放在where条件的最后，那么最后每次查询都必须要先用联合索引查询出来一部分数据，接着数据加载到内存里去，再根据where条件最后的省份、城市和性别几个字段进行过滤筛选，每次查询都得多这么一个步骤。

所以与其如此，还不如就把省份、城市和性别三个字段，放在联合索引的最左侧，这样跟其他字段组合联合索引后，让大部分的查询都可以直接通过索引树就可以把where条件指定的数据筛选出来了。

好，那么到今天为止，我们还是在分析这个案例，我们已经分析到了可以把基数较低但是频繁查询（几乎每次查询都会指定）的省份、城市和性别几个字段放到联合索引的最左侧去，此时就可以让每次查询时指定的省份、城市和性别，都直接从索引树里进行筛选。

那么联合索引中除了（province, city, sex）三个字段以外，还需要哪些其他的字段呢？其他字段该如何设计呢？是否还要设计其他的索引呢？针对这个问题，我们下一次继续分析。

**End**